

## ОТЗЫВ

официального оппонента

д.ф.-м.н. **Визильтера Юрия Валентиновича**

на диссертационную работу **Бабенко Артема Валерьевича**

«Эффективные алгоритмы поиска по большим коллекциям изображений»,  
представленную на соискание ученой степени кандидата физико-  
математических наук по специальности 05.13.11 – «Математическое и  
программное обеспечение вычислительных машин, комплексов и  
компьютерных сетей»

Работа А. В. Бабенко посвящена разработке эффективных алгоритмов визуального поиска по сходству, предназначенных для применения, в первую очередь, в современных поисковых системах. Система визуального поиска получает на вход пользовательское изображение-запрос, а в качестве ответа требуется вернуть все изображения из сети Интернет, содержащие тот же объект или сцену, что и запрос. Поскольку количество пользовательских фотографий в глобальной сети Интернет все время увеличивается, поисковым системам приходится обрабатывать постоянно растущие коллекции из миллиардов изображений. При этом необходимо обеспечить время ожидания пользователем ответа на запрос порядка сотен миллисекунд, что требует разработки эффективных алгоритмов быстрой индексации поисковых коллекций и быстрого поиска в подобных гигантских коллекциях изображений. Кроме того, в настоящее время многие поисковые приложения должны полноценно функционировать на мобильных устройствах, ограниченная оперативная память которых не вмещает поисковую коллекцию. В связи с этим возникает необходимость хранения поисковой коллекции в сжатом виде, обеспечивающем, тем не менее, эффективную реализацию процедуры поиска.

В диссертационной работе А. В. Бабенко предложен полный протокол построения системы визуального поиска, включающий как алгоритм сжатия поисковой коллекции, так и алгоритм быстрого поиска по ней. Подобная система действительно позволяет осуществлять поиск по коллекциям из миллиардов изображений за время порядка сотен миллисекунд. Таким образом, тема диссертационного исследования А. В. Бабенко не только является **актуальной**, но имеет прямое практическое применение.

В работе автором получен ряд **новых и оригинальных результатов**.

Предложены способы формирования векторных описаний изображения на основе глубоких нейронных сетей. Показано, что качество предложенных дескрипторов практически не снижается при сжатии их методом главных компонент. Предложен новый метод извлечения дескрипторов из сверточных слоев нейросети путем поэлементного суммирования, преимущество которого заключается в том, что дескрипторы могут быть вычислены для изображения любого размера и формы. Проведено экспериментальное сравнение, подтверждающее, что векторные описания, формируемые предложенными методами, позволяют достичь существенно более высокого качества поиска по сравнению с существующими методами. Также продемонстрировано, что предложенные подходы позволяют формировать векторные описания меньшей размерности, причем для качественного поиска по большинству коллекций достаточно размерности вектора признаков порядка ста. Благодаря столь компактному описанию, разработанные А.В. Бабенко алгоритмы, могут быть использованы для поиска в коллекциях большего объема, по сравнению с аналогичными методами, оперирующими описаниями большей размерности.

На основе новой модели аппроксимации векторов с помощью суммы слов (в отличие от конкатенации в модели мультиквантизации) предложено два новых подхода к сжатию векторов высокой размерности, которые, в

отличие от существующих подходов, не накладывают серьезных ограничений на распределение сжимаемых данных. Предложенный алгоритм аддитивной квантизации не накладывает ограничений на значения слов и приводит к наилучшему качеству сжатия, но требует решения сложной оптимизационной задачи, что приводит к вычислительной неэффективности при больших длинах кода. Второй предложенный алгоритм древесной квантизации накладывает некоторые ограничения на слова, что снижает количество настраиваемых параметров, зато обеспечивает высокую скорость кодирования. Условие ортогональности слов из различных словарей, не связанных по дереву, позволяет построить эффективную процедуру кодирования с использованием глобальной оптимизации методом динамического программирования. Теоретически доказано, что предложенное обобщение позволяет достигать меньшей ошибки сжатия по сравнению с известным методом мультиквантизации. Приведено большое количество экспериментальных результатов, подтверждающих преимущество предложенного подхода на практике.

Разработаны две структуры данных (ортогональный и неортогональный инвертированные мультииндексы), которые позволяют с высокой точностью и скоростью искать в сверхбольших коллекциях данных. Мультииндексы разбивают поисковое пространство на большое число областей такой структуры, которая позволяет эффективно находить наиболее перспективные области для заданного запроса. В диссертации описана эффективная процедура нахождения ближайших регионов для конкретного запроса и доказаны оценки асимптотической алгоритмической сложности такой процедуры. Экспериментально продемонстрировано, что предложенные автором структуры данных позволяют искать в миллиардных коллекциях векторов за несколько миллисекунд с такой точностью, для достижения которой всем известным аналогам требуется более ста миллисекунд.

**Практическая значимость** работы заключается в том, что на основе предложенных методов разработана целостная система визуального поиска, которая успешно используется в поиске по изображениям компании ООО «Яндекс».

В качестве **замечаний** по содержанию и оформлению данной работы можно указать следующие:

1. Использование в автореферате диссертации конструкций типа «автор предлагает», «автор описывает» и т.п. выглядит несколько странным, ведь это *автореферат*. Более предпочтительным представляется использование безличных соответствующих конструкций «предложено», «описано» и т.п.

2. В подходе к сжатию векторов, основанном на аддитивной квантизации, представляется интересным рассмотреть возможность рекурсивного кодирования новыми словарями остатков, образованных разностью исходных и уже закодированных предыдущими словарями данных.

3. В работе упоминаются методы снижения размерности данных и поиска в коллекциях данных, основанные на бинарном семантическом хешировании, однако сравнение с такими методами по степени сжатия, точности и скорости поиска не проводится. Желательно было бы такое сравнение провести, поскольку последние результаты в данной области, где хешированию подвергаются признаковые описания, сформированные глубокими нейронными сетями, выглядят достаточно конкурентоспособными, в частности, в задачах поиска в коллекциях изображений лиц.

4. Современный подход к формированию целевой функции при решении задач снижения размерности дескриптора в задачах поиска по сходству предполагает оптимизацию не точности воспроизведения исходных кодируемых данных, а непосредственно точности поиска в коллекциях. При этом, как было, в частности, показано в ряде работ по семантическому

хешированию, кодирование со снижением размерности может не только не снижать точность поиска, но иногда даже повышать ее.

Указанные замечания не снижают научной и практической значимости диссертационной работы. Работа производит целостное впечатление, написана понятно и выполнена на высоком научном уровне, содержит решение актуальной задачи поиска данных в больших коллекциях изображений. Описанные выше оригинальные результаты работы не имеют аналогов в известных оппоненту публикациях, т.е. являются новыми научными результатами. **Содержание диссертации** в достаточной степени опубликовано в научной печати – в ряде публикаций, из которых 6 в научных изданиях, рекомендованных ВАК. Основные результаты докладывались на ведущих российских и международных конференциях. Имеются публикации, индексируемые в Scopus и Web of Science. **Научная достоверность и обоснованность** полученных результатов определяются корректностью математических доказательств свойств предложенных методов, согласованностью выводов и данных детально описанных воспроизводимых экспериментов, а также публично доступным программным кодом.

**Автореферат** достаточно полно и правильно отражает содержание диссертации. Тематика и содержание диссертационного исследования соответствуют специальности 05.13.11 «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

В целом диссертационная работа «Эффективные алгоритмы поиска по большим коллекциям изображений» представляет собой законченное научное исследование, выполненное на высоком научном уровне на актуальную тему, содержит новые научные результаты, имеющие существенное значение для рассматриваемой предметной области, и отвечает всем требованиям, предъявляемым ВАК к кандидатским диссертациям (п. 9 «Положения о присуждении ученых степеней», утвержденного Постановлением Правительства Российской Федерации от 24.09.2013, №842),

а ее автор – Бабенко Артем Валерьевич – заслуживает присуждения ему учёной степени кандидата физико-математических наук по специальности 05.13.11 «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

Официальный оппонент,  
начальник подразделения  
ФГУП «ГосНИИАС»,  
д.ф.-м.н., профессор РАН

Визильтер Юрий Валентинович

«2» мая 2017 г.

Подпись Визильтера Ю. В. заверяю.

Ученый секретарь ФГУП «ГосНИИАС»,  
д.т.н.



Мужичек Сергей Михайлович

«\_\_» \_\_\_\_\_ 2017 г.

Визильтер Юрий Валентинович – доктор физико-математических наук, старший научный сотрудник, профессор РАН, начальник подразделения "Системы интеллектуального анализа данных, технического зрения, улучшенного и синтезированного видения" Федерального государственного унитарного предприятия "Государственный научно-исследовательский институт авиационных систем" (ФГУП "ГосНИИАС").

г. Москва, ул. Викторенко, 7, тел.: (499) 157-94-98, e-mail: viz@gosniias.ru.